International Journal of Innovative Computing, Information and Control
Volume **19**, Number **5**, October **2023**

ICIC International ©2023 ISSN 1349-4198
pp. **1361–1375**

DAS-VSP NOISE ELIMINATION BASED ON THE DILATED PYRAMID ATTENTION NETWORK

Tie Zhong[1], Haoliang Chen[1], Shiqi Dong[2] and Jianpo Li[3,*]

[1]School of Electrical Engineering
[3]School of Computer Science
Northeast Electric Power University
No. 169, Changchun Road, Jilin 132012, P. R. China
zht@neepu.edu.cn; 1017128732@qq.com
*Corresponding author: lijianpo@neepu.edu.cn

[2]Key Laboratory of Geotechnical Mechanics and Engineering of the Ministry of Water Resources
Changjiang River Scientific Research Institute
No. 23, Huangpu Road, Wuhan 430010, P. R. China
dsq1994@126.com

Received December 2022; revised March 2023

Abstract. *Distributed acoustic sensing (DAS) is widely used in seismic data acquisition due to its low cost and high acquisition density. However, the collected DAS records are often accompanied by various interferences, which can impact the recognition of effective information and subsequent processing. Traditional noise elimination methods have limitations and cannot fulfill the requirements. Therefore, deep learning-based methods have been introduced to solve the noise elimination problem of DAS records. However, these networks may have deficiencies in their structures, leading to a degradation in their denoising ability. To achieve denoising capabilities, we propose a multi-scale network named the dilated pyramid attention network (DPA-Net) to tackle complex DAS background noise. The DPA-Net uses pyramid-shaped modules to extract features at different scales, employs a serration dilated convolution in the pyramid module to suppress the checkerboard effect and increase the receptive field, and uses a multi-scale module to optimize the extraction of coarse features at low resolutions. Additionally, an attention module is employed to enhance the effective features, thus improving the denoising performance. Experiments on both synthetic and field records have verified that DPA-Net can effectively suppress complex DAS background noise and recover weak reflection signals.*
Keywords: Background noise suppression, Convolutional neural network (CNN), Distributed acoustic sensing (DAS), Vertical seismic profile (VSP)

1. **Introduction.** Distributed acoustic sensing (DAS) utilizes changes in the phase information of scattered light in optical fibers to detect slight deformations in the ground caused by seismic waves [1]. Compared to traditional geophone arrays, DAS has advantages in terms of arrangement and detection density [2]. Recently, it has been considered an emerging information acquisition technology and has been applied in vertical seismic profiling (VSP). However, sensitive scattered signals in DAS recording systems are easily affected by complex background noise [3]. Additionally, the DAS background noise is primarily composed of instrument noise and coupled noise, some of which are not present in conventional seismic records, such as time-varying optical noise and horizontal noise [4,5]. This implies that conventional denoising methods may not be effective in dealing with the different properties of DAS background noise, which differ from conventional seismic

DOI: 10.24507/ijicic.19.05.1361

random noise. Improving the effective signal and reducing background noise is a crucial step in DAS data processing [6], and as a result, the denoising of DAS data has received growing attention in the seismic exploration industry.

In general, conventional denoising methods can be divided into four categories: classical methods, time-frequency-based methods, multiscale-decomposition methods, and sparsity-based methods. Classical methods, such as band-pass filtering (BPF) [7,8], median filtering [9,10], and Wiener filtering [11-13], can separate the desired signals from the noisy records by exploiting differences in physical properties or frequency components. However, these methods have simple denoising principles and often underperform when dealing with complex seismic data [14]. For instance, Wiener filtering has limited effectiveness in eliminating non-stationary noise [15]. Time-frequency based methods, such as S-transform [16] and time-frequency peak filtering [17,18], have been introduced as an improvement over classical methods, using the properties of signals within the time-frequency domain to separate signal and noise components. However, determining the filtering parameters for these methods can be challenging, as effective signals and background noise are mixed in the time-frequency domain, and inappropriate parameters can result in signal loss or residual noise [19]. To address these limitations, multiscale-decomposition methods, such as empirical modal decomposition (EMD) [20,21], ensemble EMD [22], and variational mode decomposition [23], have been proposed. These methods decompose the analyzed records into intrinsic modes, recovering desired signals by keeping signal-dominant modes and discarding noise-dominant modes. However, DAS records often contain high-frequency background noise, low-frequency horizontal noise, and low-frequency and high-amplitude time-varying optical noise superimposed on the effective signals, leading to mixing of modes representing effective signals and noise in the same intrinsic mode functions (IMFs). This results in limited noise removal even when a part of the IMF that completely represents noise is removed, as some noise modes and effective signal modes are mixed together and cannot be separated. Mode-mixing, which occurs when different IMFs have the same or similar time scales, further exacerbates the situation and seriously affects the denoising performance of multiscale-decomposition methods, especially in the presence of aliasing noise [24,25]. Sparsity-based methods, such as wavelet transform [26], shearlet [27], seislet [28], curvelet [29], and dictionary learning methods [30], are also widely used for denoising. These methods rely on converting noisy data into the sparse domain and separating signal and noise components using threshold functions to eliminate noise. While these methods are often effective for seismic data, as it can be considered sparse, determining the threshold functions is typically empirical and done manually, hindering their widespread use. Other methods, such as singular value decomposition [31], principal component analysis [32], and diffusion filtering [33], have been applied to seismic data processing, but have not met the requirements for complex DAS background noise elimination according to our trials. With the rapid of denoising theory, networks based on deep learning have been extensively discussed and gradually applied in the seismic data processing [34,35]. Some successful applications are reported in the field of inversion [36], waveform classification [37], and interpolation [38]. Meanwhile, the convolutional neural network (CNN), as the important content of deep learning, has also been introduced in seismic noise attenuation. The basic principle for CNN-based networks is to establish a non-linear mapping relation between the desired signals and noisy data through the training process [39]. Here, some typical methods, such as feedforward denoising convolutional neural networks (DnCNNs) [40,41] and generative adversarial networks (GAN) [42], have employed and represented better performance than the conventional methods [43]. The excellent performance is mainly determined by the following two factors: network architecture and training dataset. For the network architecture, the conventional CNNs always

have a relatively simple architecture, such as the backbone of DnCNN. There are no effective feature interaction operations in the network, thereby hindering further improvement in feature exaction and denoising performance [44-47]. For the training dataset, it is unable to capture the desired signals from the field data. Thus, how to generate a high-quality signal training dataset is a critical issue that should be solved. As we know, the inappropriate training data will result in amplitude loss or property changes for the recovered signals [48]. Therefore, we need to design an effective network and construct a generalized training dataset to guarantee the demands of DAS data processing.

In order to effectively suppress noise, a novel network called the Dilated Pyramid Attention Network (DPA-Net) is proposed in this paper. In general, DPA-Net incorporates a pyramid-like structure and a multi-scale attention module to achieve an optimal representation of potential features. The pyramid-like backbone extracts coarse information from DAS records, while shallow information is introduced through skip connections to enhance effective information. Serration dilated convolutional layers are used to increase the receptive field and prevent the checkerboard effect. The attention module reinforces edge information and further improves denoising capability. The network is trained using a high-quality training dataset consisting of both pure synthetic signals and real DAS background noise to ensure efficient network training. DPA-Net is tested using both synthetic data and field records, and its performance is compared with that of commonly used methods.

The rest of this paper is organized as follows. Section 2 provides a detailed explanation of the network and its components, including the construction of the training set, experimental environment, and network parameters. Section 3 presents the results of comparative experiments. Section 4 concludes the paper.

2. **Method.** The network structure of DPA-Net is shown in Figure 1. In general, DPA-Net is comprised of a pyramid module, a multi-scale module, and an attention module based on serration dilated convolution (SDCONV), and has a multi-scale architecture. Specifically, the network backbone is a codec process that uses up-sampling and down-sampling to obtain rich, multi-resolution features. The multi-scale module, combined with skip connections, is then utilized to extract deep features and fuse shallow information. SDCONV is applied to enlarging the receptive field and improving denoising efficiency. The attention module is used in the decoding process to further improve the feature

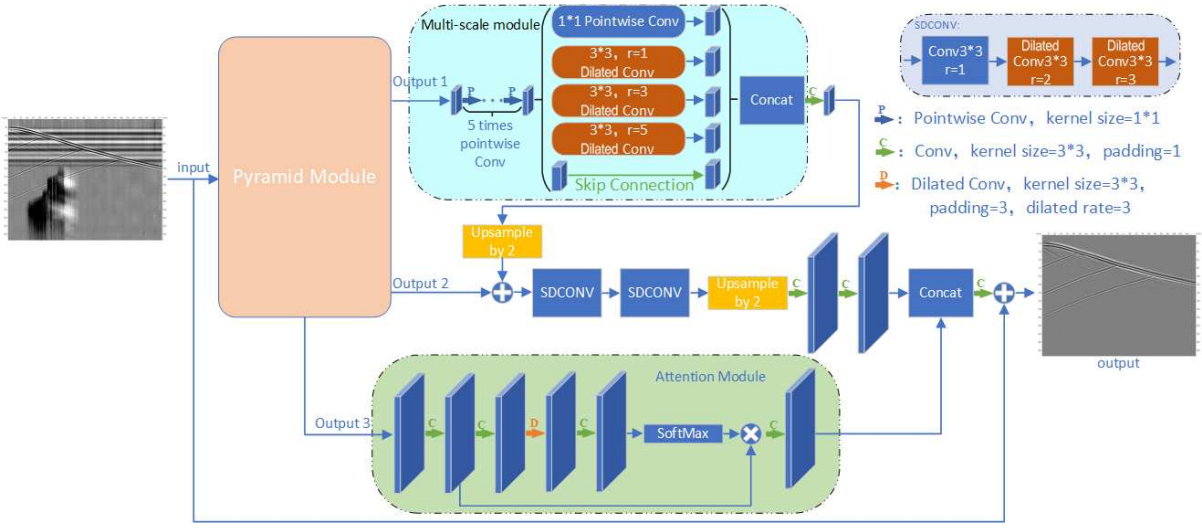


FIGURE 1. DPA-Net architecture

extraction capability by fusing potential features and enhancing the extraction of edge features. The specific components of the network are described as follows.

2.1. **Pyramid module.** The structure of the pyramid module in DPA-Net is depicted in Figure 2. It can be considered as an encoding module. To begin, we use five convolutional layers to extract the potential features in the input data. In this process, we employ a convolutional layer with a stride of 2 to perform the down-sampling operation on the original input data. Then, two SDCONVs are used to further extract informative features and fuse them with shallow features after down-sampling. We repeat this process to extract the coarse features from low-resolution data. This allows us to capture features at different resolutions and use them as inputs for the following blocks.

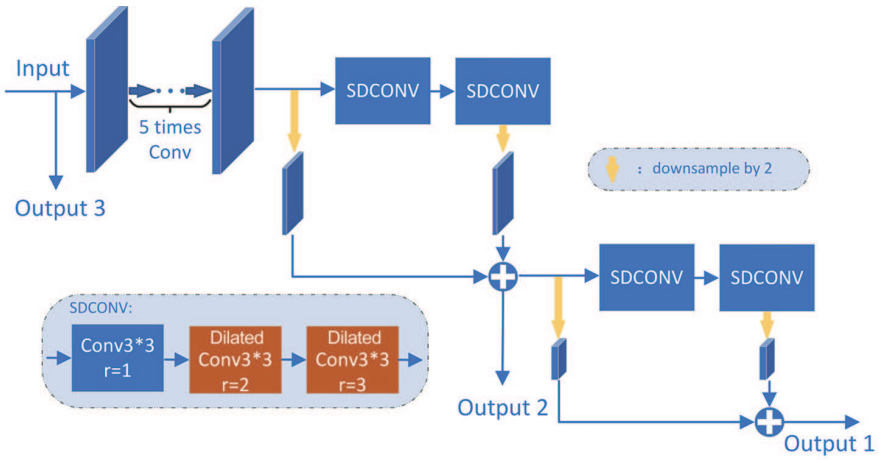


FIGURE 2. Pyramid module architecture

2.2. **Attention module.** In general, the DAS signals are always seriously affected by complex background noise, making it a challenging task to effectively recover the desired signals. To enhance the feature extraction capability and reinforce the denoising performance in the face of weak upgoing signals, we use a spatial attention module. Specifically, the attention module uses five convolutional layers for initial feature extraction. To improve the convolution efficiency, dilated convolution is interspersed between them. The output of the second convolutional layer is then activated by the SoftMax function and multiplied with it for self-enhancement. The SoftMax function is a commonly used activation function for self-enhancement and its formula is shown below:

$$SoftMax(Z_i) = \frac{\exp(Z_i)}{\sum_j \exp(Z_j)} \tag{1}$$

where Z is the element in the attention map, i is the current element label, and j is the number of all elements. The denominator normalizes the function output to a probability value in the range $[0, 1]$. Because the numerator uses the exponential function, the dominant features that are most easily extracted receive higher probability values. In the denoising of the DAS record, this will make the edge information in the attention map more prominent and increase the effectiveness of the attention module. Finally, the output of this module is fused with the backbone of DPA-Net in the channel dimension (by concatenation) to enhance the informative features. In general, this module reinforces the effective features in DAS records, making low-amplitude reflection signals more easily recognizable by the network.

2.3. **Serration dilated convolution (SDCONV).** Dilated convolution has the effect of increasing the receptive field. Its principle is to interpolate zeros to the convolution kernel and obtain a larger receptive field without increasing the computational cost. In this paper, dilated convolutional layers with different dilated rates are used to enlarge the receptive field. However, continuous use of dilated convolutions will cause the checkerboard effect, resulting in information loss and serious degradation of the network performance. Therefore, the same dilated rates should not be used continuously when designing dilated convolution, and the principle of hybrid dilated convolution (HDC) should be followed [49]. For a series of dilated convolutions whose convolution kernel size is $K \times K$ and dilated rates are $[r_1, r_2, \ldots, r_i, \ldots, r_n]$, we define the maximum distance M_i between two non-zero elements, shown as follows:

$$M_i = \max[M_{i+1} - 2r_i, 2r_i - M_{i+1}, r_i] \tag{2}$$

where M_i is the maximum interval between two non-zero elements in the i layer; r_i is the number of zeros in the i layer. If $M_2 \leq K$, then this series of dilated convolution can completely cover the underlying feature region. After trying various dilated rates, we found that the convolutional layers with dilation = 1 can extract features better. In this paper, we experiment with different convolution kernels. Compared to the 3×3 convolution kernel, the 7×7 convolution kernel has a slight degradation in the SNR, while the 5×5 convolution kernel has a slight improvement of 0.01 dB, but its training and testing time increases significantly. After careful consideration, we chose the 3×3 convolution kernel. The dilated rates affect the cumulative receptive field, and a larger cumulative receptive field focuses more on global features and less on detailed feature extraction. In the denoising of DAS records, we are more concerned with local effective signal retention, so a larger receptive field can affect effective signal retention, especially for low-amplitude signals. When $r = [1, 2, 3]$, the cumulative receptive field is 13×13, but when $r = [1, 2, 5]$, the cumulative receptive field is 17×17. With a minimum feature map size of only 16×16, using $r = [1, 2, 5]$ results in a more global feature extraction and reduces the ability to extract detailed features. Our experiments show that using $r = [1, 2, 5]$ results in a 0.17 dB lower signal-to-noise ratio compared to using $r = [1, 2, 3]$. After careful consideration, we choose the convolution set with $K = 3$ and $r = [1, 2, 3]$, referred to as SDCONV, for use in this paper.

2.4. **Seismic denoising principle.** In seismic exploration, the noisy data can be regarded as the superposition of effective signals and various complex noises, shown as follows:

$$y = s + n \tag{3}$$

where y represents the noisy DAS-VSP record, s and n denote the effective signal and background noise, respectively. Subsequently, we construct the denoising network and obtain the high-dimensional mapping relationship between noisy records and effective signals after training. Moreover, the estimated signal can be represented as

$$\hat{s} = F(y; \zeta) \tag{4}$$

where $\hat{s}$ is an estimate of effective signal, F represents this high-dimensional mapping relationship. In addition, $\zeta = \{w, k\}$ is the parameter set, while w and k represent the weight and bias, respectively. From the basis, we use the loss function to denote the estimation errors:

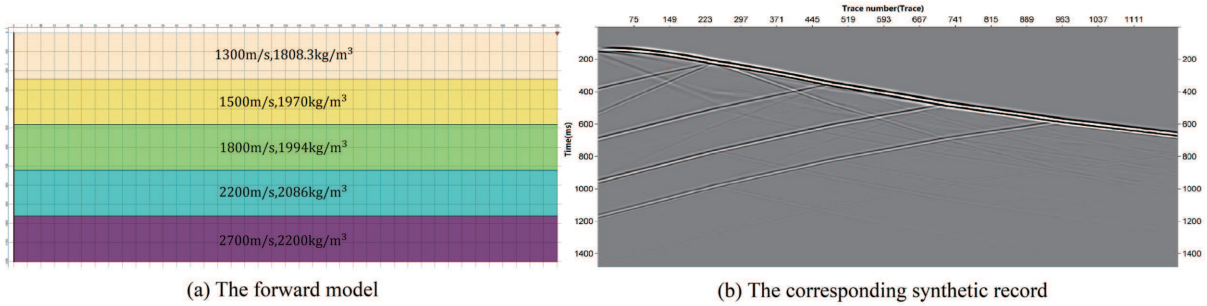
$$L(\zeta) = \frac{1}{N} \sum_{i=1}^{N} \|F(y_i; \zeta) - S_i\|^2 \tag{5}$$

where y_i denotes the noisy records and S_i denotes the effective signal patches. It can be seen that the network learns the effective signal in the training set and optimizes the parameters gradually through iterations to achieve the optimal signal estimation results. Thus, the parameter set ζ_{out}, minimized the loss function, is chosen as the optimal parameter. Then, the denoised signals can be expressed:

$$\hat{s}_{out} = F(y; \zeta_{out}) \tag{6}$$

2.5. **Construction of the training set.** As we know, the excellent performance of deep learning methods largely depends on the diversity and generalization of the data sets. In this study, the network training data set consists of an effective signal set and a field noise set. In previous studies, the pure signal set was always based on some simulations of seismic wavelets. However, the generated effective signals were relatively simple and had a difference in properties with field data. This resulted in the trained network models performing poorly in the recovery of weak reflection signals in DAS records. Therefore, the forward modeling method was applied to generating high-quality DAS data, ensuring the training accuracy of the trained models.

For obtaining effective signals, we use previously acquired profile data as a reference to construct a forward stratigraphic model. Wavelets with different fundamental frequencies are used as input to the wave equation, and the synthetic record is generated based on the stratigraphic model. In this study, 20 forward models were generated, producing 150 synthetic records. The diversity of reflection events ensures the generalization ability of the trained model. Figure 3(a) shows the constructed forward modeling model with wave velocities of 1300 m/s, 1500 m/s, 1800 m/s, 2200 m/s, and 2700 m/s. The red triangle in the upper right corner represents the location of the artificial source, and the black vertical line on the left represents the DAS downhole receiving line. Figure 3(b) shows the corresponding synthetic record based on the forward model. Synthetic records are then cut into 64×64 small signal patches, and 20,000 patches are randomly selected to compose the effective signal set. The real noise set is obtained from actual DAS records obtained without artificial source excitation during an exploration in the northeast of China. The real noise is divided into three categories: fading noise, horizontal noise, and abnormal interference. The noise data is separated into 64×64 patches, and 30,000 patches are randomly selected to compose the noise dataset. The diversity of noise types guarantees the accuracy of the training. Some typical signal and noise patches are shown in Figure 4. During the training process, the effective signal set is randomly combined with the real noise set, and the effective signal patches and noisy data are used as inputs for the network.



(a) The forward model (b) The corresponding synthetic record

FIGURE 3. Construction of clean synthetic record

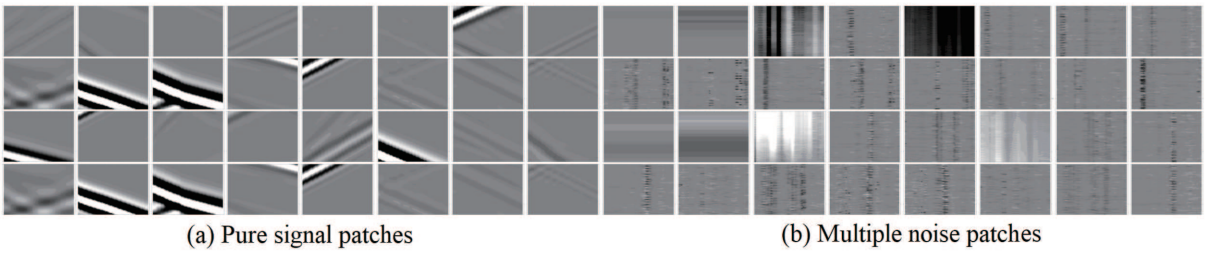


FIGURE 4. Samples of the training dataset

2.6. **Training process and experimental environment.** Deep learning requires high hardware conditions, and the experimental environment in this paper consists of an Intel(R) Core(TM) i5-9400f, an NVIDIA Geforce RTX 2060 Super, and 16 GB of RAM. For comparison purposes, all experimental methods use the same training set and experimental environment. Specifically, the ADAM optimizer is used with learning rates of $\left[10^{-4}, 10^{-5}, 10^{-6}\right]$ to optimize the network. The batch size is set to 32, and the training process consists of 60 epochs. The network training parameters for the deep learning methods are shown in Table 1.

TABLE 1. Network architecture parameters

Hyper-parameter	DnCNN	U-Net	DPA-Net
Optimizer	ADAM	ADAM	ADAM
Patch size	64×64	64×64	64×64
Batch size	32	32	32
Epoch	60	60	60
Learning rate range	$\left[10^{-4}, 10^{-5}, 10^{-6}\right]$	$\left[10^{-4}, 10^{-5}, 10^{-6}\right]$	$\left[10^{-4}, 10^{-5}, 10^{-6}\right]$
Input channels	1	1	1
Layers	17	19	33
Convolution kernel size	$3 \times 3 \times 64$	$3 \times 3 \times 64$	$3 \times 3 \times 64$ or $1 \times 1 \times 64$

3. Results.

3.1. **Synthetic data processing.** To test the results of the network, we generate a pure composite record using the forward modeling method described above. It should be noted that the stratigraphic data used in this forward simulation is different from the 150 stratigraphic data in the training set to ensure the independence of the test. The model consists of four layers with wave velocities of 1500 m/s, 1800 m/s, 2200 m/s, and 2700 m/s. The forward model is activated using a Ricker wavelet with a dominant frequency of 40 Hz. To test the effectiveness of the network in removing real noise, we add field noise data, which was not included in the training dataset, as shown in the bottom subplot of Figure 5(a). The noise is then added to the pure synthetic record to obtain the noisy synthetic record, as shown in the top subplot of Figure 5(a).

In this paper, BPF and some CNN-based networks are selected as competing methods. Specifically, we set the passband range of the BPF to [8-46 Hz] by analyzing the power spectrum characteristics of the synthetic recordings. Additionally, CNN-based methods, including DnCNN and U-Net, are applied to the noisy synthetic record. To ensure the fairness of the comparison, the CNN-based methods use the same configuration and training set.

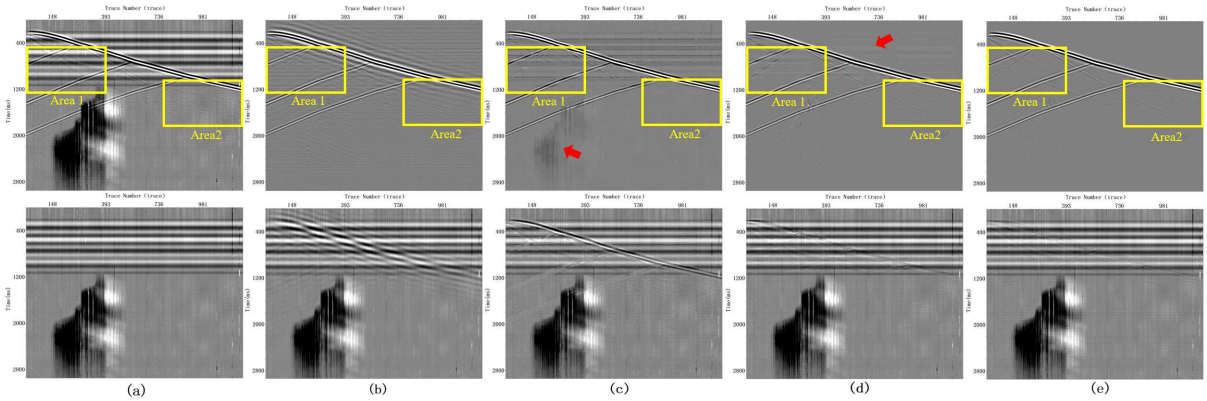


FIGURE 5. Processing results of synthetic record: (a) Noisy records (top subplot) and added noise (bottom subplot); (b)-(e) processing results of BPF, DnCNN, U-Net, DPA-Net (top subplot) and filtered noise (bottom subplot)

The processing results are shown in Figure 5. Due to the nature of BPF, it can only eliminate the noise components outside the passband, resulting in limited separation of the signal and noise within the passband, as shown in Figure 5(b). As a result, obvious signal leakages can be seen in the filtered noise, indicating negative impacts on the preservation of signal amplitude. Conversely, CNN-based methods provide better results compared to BPF. Notably, DnCNN (Figure 5(c)) is capable of suppressing fading noise, but its denoising performance degrades when faced with complex noise such as residual time-variant optical noise (marked by red arrows). Additionally, conspicuous signal leakage shows its weakness in signal recovery. Although U-Net (Figure 5(d)) can eliminate time-variant optical noise, it still has residual horizontal noise (marked by red arrows), affecting the denoising result. Overall, DPA-Net exhibits excellent noise elimination ability, as the result shows a clean background and the recovery of some small signals in deep strata. To make more detailed comparisons, two areas of interest, marked as yellow blocks in Figure 5, are enlarged, and the corresponding comparison results are shown in Figure 6.

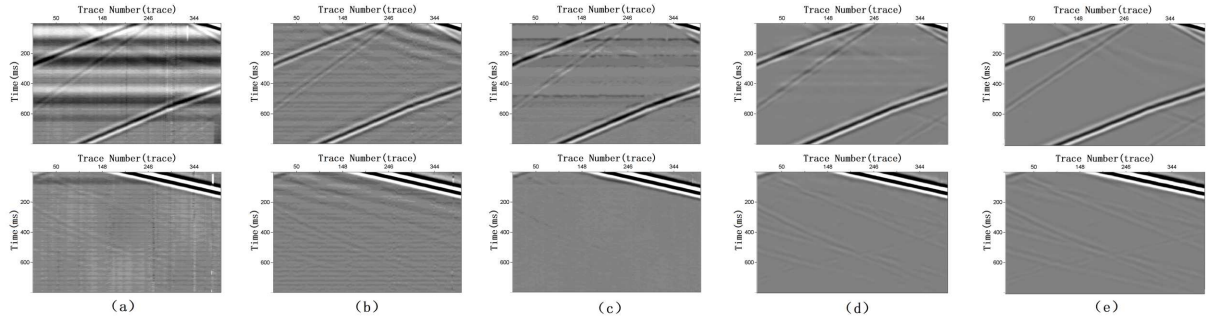


FIGURE 6. Enlarged comparisons of the synthetic results: (a)-(e) Noisy record, results of BPF, DnCNN, U-Net and DPA-Net (Area 1 (top) and Area 2 (bottom))

Notably, BPF (Figure 6(b)) has limited effects in eliminating complex DAS background noise and restoring fake signals. Although the results of CNN-based methods, DnCNN and U-Net, depicted in Figures 6(c) and 6(d), are better, they fail to completely attenuate the background noise and residual noise hinders the recognition of effective signals. In contrast, DPA-Net outperforms the competing methods in terms of signal recovery and noise

elimination. As shown in Figure 6(e), weak upgoing signals are reconstructed, verifying the effectiveness of DPA-Net in processing complex DAS data.

In order to compare the spectral information, an F-K spectrum analysis was performed, and the results are shown in Figure 7. We processed the seismic records shown in Figure 5(a), where an obvious aliasing phenomenon (Figure 7(a)) can be seen between the pure signals and field DAS noise data. It is clear that BPF (Figure 7(b)) can only eliminate the out-of-band noise components and fails to separate the signals within the passband. Similar results can be obtained by CNN-based methods, which provide better denoising performance. As shown in Figures 7(c) to 7(e), the recovered signals all have similar properties to the pure signals in the F-K domain. However, both DnCNN and U-Net still need improvement in signal-amplitude preservation as signal-leakage components can be observed in the filtered noise results. Compared to other methods, DPA-Net (Figure 7(e)) is more capable in signal recovery and noise elimination, as marked by the red blocks.

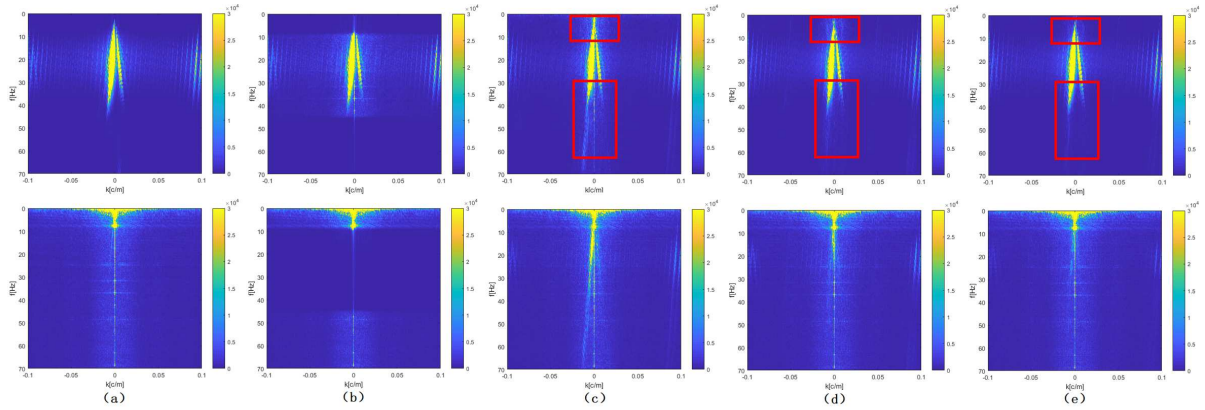


FIGURE 7. F-K spectrum for denoising results: (a) F-K spectrum for pure signals (top subplot) and added noise (bottom subplot); (b)-(e) results of BPF, DnCNN, U-Net and DPA-Net (top subplot) and filtered noise (bottom subplot)

In addition, we conduct a quantitative analysis to investigate the denoising performance. In signal processing, the signal-to-noise ratio (SNR) and root mean square error (RMSE) are two important indicators for measuring the denoising effect, and their corresponding definition equations are as follows:

$$SNR = 10 \lg \left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{m} S_{ij}^2}{\sum_{i=1}^{n} \sum_{j=1}^{m} (N_{ij} - S_{ij})^2} \right) \tag{7}$$

$$RMSE = \sqrt{\frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} (N_{ij} - S_{ij})^2} \tag{8}$$

where S represents the pure record, N is the denoising result. In addition, i and j denote the dimensions of the synthesized record. In order to investigate the denoising performance, synthetic records with different signal-to-noise ratios (SNRs) were processed, and the improved SNR and root mean square error (RMSE) results are shown in Table 2. The results indicate that the bandwidth-preserving filter (BPF) can only improve the SNR to a limited extent, with an improved SNR of only 2.34 dB for a 0 dB record, demonstrating its limited ability in noise attenuation. In contrast, the convolutional neural network (CNN)-based methods exhibit excellent denoising performance. It is shown that the

TABLE 2. The comparisons of SNR and RMSE for different attenuation algorithms

Synthetic	BPF		DnCNN		U-Net		DPA-Net	
record/dB	SNR/dB	RMSE	SNR/dB	RMSE	SNR/dB	RMSE	SNR/dB	RMSE
0	2.3484	0.3094	12.0931	0.2445	19.5480	0.1036	20.1064	0.0972
−2	1.8996	0.3861	11.0459	0.2758	17.6837	0.1285	18.4257	0.1179
−4	1.2715	0.4854	9.7947	0.3186	15.7107	0.1612	16.6384	0.1449
−6	0.4315	0.6105	8.3238	0.3774	13.6759	0.2038	14.4390	0.1866
−8	−0.6361	0.7681	6.6314	0.4585	11.6237	0.2581	12.4908	0.2336
−10	−1.9245	0.9666	4.7369	0.5703	9.5948	0.3260	10.5569	0.2918

DnCNN and U-Net methods can achieve an average improvement of 13.77 dB and 19.64 dB, respectively. Compared to the competing methods, the DPA-Net has the best denoising capability, with a nearly 20 dB improvement in SNR for −10 dB data. Similar results were also obtained for the RMSE. Therefore, the improved SNR and RMSE results demonstrate that the DPA-Net is effective in complex DAS background noise elimination, especially for low-SNR data.

Meanwhile, the training and testing speeds of neural networks are equally important. Table 3 summarizes the testing time, training time, and average SNR increment of the network and competing methods. BPF has the shortest testing time and does not require training, but its average SNR increment is too low to achieve effective denoising. The training time and testing time of DnCNN and U-Net are roughly the same, with DnCNN having a good denoising effect. The test time of DPA-Net is roughly the same as that of U-Net, and its training time is slightly longer. However, considering its high SNR increment with excellent signal preservation and noise suppression capabilities, the computing cost is acceptable. Therefore, it can be concluded safely that DPA-Net has the advantage in processing DAS data and eliminating complex background noise.

TABLE 3. Processing time and improved SNR for different denoising methods

Original record/dB	BPF	DnCNN	U-Net	DPA-Net
Processing time (s)	0.72	1.16	1.17	1.18
Training time (hour)	–	2.4	2.7	3.4
Averaged improved SNR (dB)	5.7771	13.77	19.64	20.44

3.2. **Field data processing.** In this paper, the denoising performance of DPA-Net in field DAS data is also investigated. A field DAS record acquired in Northeast China is processed and shown in Figure 8(a). Notably, the record consists of 1368 trace records with a sampling interval of 4 μs. It is evident that the reflection events are heavily contaminated by complex DAS background noise, such as coupled noise and time-variant noise (marked by the blue blocks). We applied the aforementioned methods to processing the field data and the denoising results are shown in Figure 8.

Owing to its denoising principle, BPF (Figure 8(b)) has residual noise that shares the same frequency bands as the desired signals. Furthermore, the recovered signals appear to have changes in frequency components as the events become broader. While DnCNN (Figure 8(c)) can effectively suppress the noise, the residual noise still severely affects the data quality. On the other hand, U-Net (Figure 8(d)) is significantly better than DnCNN in suppressing complex noise, and the recovered signals are continuous. However, the residual noise, such as horizontal noise (marked by red arrows), can still be observed. In

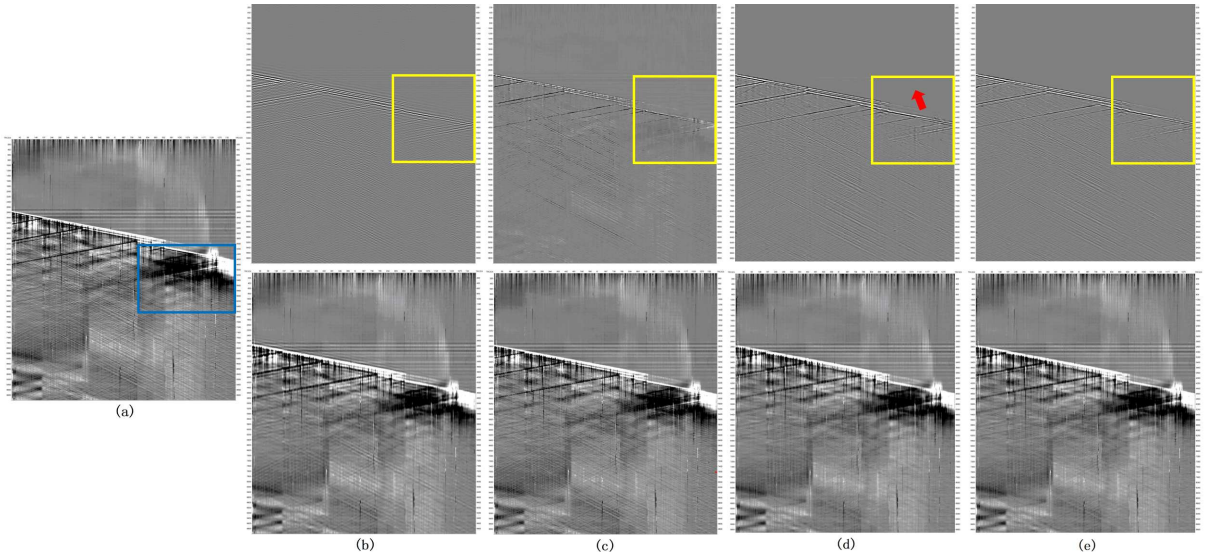


FIGURE 8. (color online) Processing results of field data: (a) Field DAS record; (b)-(e) results (top subplot) and filtered noise (bottom subplot) of BPF, DnCNN, U-Net and DPA-Net

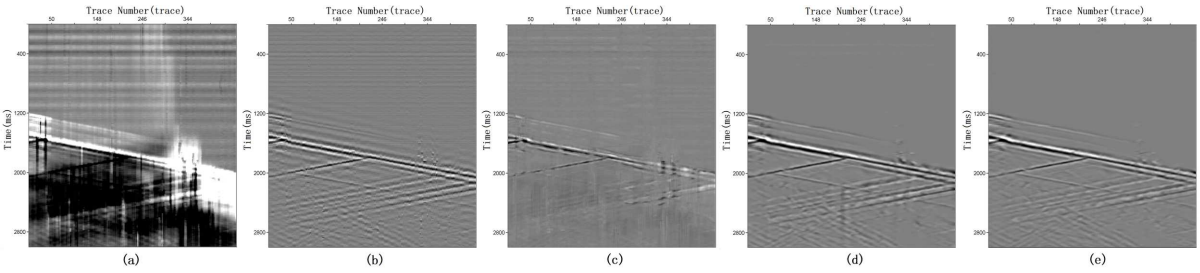


FIGURE 9. Enlarged comparisons of field data: (a) Field DAS record; (b)-(e) enlarged denoising results of BPF, DnCNN, U-Net, and DPA-Net

contrast, DPA-Net (Figure 8(e)) demonstrates excellent results in both signal recovery and noise elimination. For clarity, the areas marked by yellow blocks in Figure 8 have been enlarged, and the corresponding results are shown in Figure 9. Similar to the results in Figure 8, BPF (Figure 9(b)) fails to effectively eliminate complex noise. Among the results of the CNN-based methods, as shown in Figures 9(c) to 9(e), only DPA-Net can restore the desired signals and significantly improve the quality of the DAS data. These experimental results verify the effectiveness of DPA-Net in eliminating complex DAS background noise.

4. **Conclusion.** In this paper, we propose a multi-scale denoising network based on a dilation pyramid and attention mechanism to reduce the complex background noise in DAS-VSP records. The network employs a pyramid module for multi-resolution feature extraction and uses serrated dilation convolutions in the pyramid module to avoid the checkerboard effect and improve feature extraction by increasing the receptive field. Additionally, a multi-scale module is designed to extract features at different scales, and the attention module enhances the primary features, thereby enhancing the denoising capability. To train the network, we create a high-quality training dataset using synthetic records and field noise data. The performance of DPA-Net and other competing methods are evaluated by processing both synthetic data and a field record. The results show that DPA-Net is effective in suppressing complex noise and reconstructing the desired signals,

even in deep strata with weak signals. Thus, DPA-Net has research significance and reference value in processing DAS records. However, the network has some limitations, such as mediocre performance on low signal-to-noise ratio DAS records. The denoising of convolutional neural networks is mainly based on feature extraction, and thus, it is a future research direction to seek a network with stronger feature extraction ability. With the increasing demand for energy and the development of seismic exploration in challenging environments, such as mountains and oceans, there will be new data types, which will challenge existing denoising methods. Hence, it is important to design new denoising methods to meet the needs of denoising for these new data types.

REFERENCES

[1] K. Spikes, N. Tisato, T. Hess et al., Comparison of geophone and surface-deployed distributed acoustic sensing seismic data, *Geophysics*, vol.84, no.2, pp.A25-A29, 2019.

[2] T. Daley, B. Freifeld, J. Ajo-Franklin et al., Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring, *The Leading Edge*, vol.32, no.6, pp.699-706, 2013.

[3] G. Binder, A. Titov, Y. Liu et al., Modeling the seismic response of individual hydraulic fracturing stages observed in a time-lapse distributed acoustic sensing vertical seismic profiling survey, *Geophysics*, vol.85, no.4, pp.T225-T235, 2020.

[4] G. Yu, J. Greer, Q. Zhang et al., Walkaway VSP using multimode optical fibers in a hybrid wireline, *The Leading Edge*, vol.35, no.7, pp.615-619, 2016.

[5] S. Wang, Y. Li and Y. Zhao, Attribute-guided target data separation network for DAS VSP data, *IEEE Transactions on Geoscience and Remote Sensing*, vol.60, 2022.

[6] A. Mateeva, J. Lopez, H. Potters et al., Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling, *Geophysical Prospecting*, vol.62, no.4, pp.679-692, 2014.

[7] W. Liu, S. Cao and Y. Chen, Seismic time-frequency analysis via empirical wavelet transform, *IEEE Geoscience and Remote Sensing Letters*, vol.13, no.1, pp.28-32, 2016.

[8] X. Dong, H. Jiang, S. Zheng, Y. Li and B. Yang, Signal-to-noise ratio enhancement for 3C downhole microseismic data based on the 3D shearlet transform and improved back-propagation neural networks, *Geophysics*, vol.84, no.4, pp.V245-V254, 2019.

[9] R. Czerwinski, D. Jones and W. D. O'Brien, Ultrasound speckle reduction by directional median filtering, *International Conference on Image Processing*, pp.358-361, 1995.

[10] R. Vijayarajan and S. Muttan, Clustering performance analysis of FCM algorithm on iterative relaxed median filtered medical images, *The 3rd International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom 2011)*, Bangalore, pp.152-157, 2011.

[11] W. Gardner, Cyclic Wiener filtering: Theory and method, *IEEE Transactions on Communications*, vol.41, no.1, pp.151-163, 1993.

[12] B. Reddy and V. Jayaraman, Application of Wiener filter making signals orthogonal, *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, Vellore, India, pp.1-6, 2019.

[13] E. Robinson and S. Treitel, Principles of digital Wiener filtering, *Geophysical Prospecting*, no.15, pp.311-332, 2006.

[14] X. Dong, Y. Li, T. Zhong, N. Wu and H. Wang, Random and coherent noise suppression in DAS-VSP data by using a supervised deep learning method, *IEEE Geoscience and Remote Sensing Letters*, vol.19, pp.1-5, 8001605, doi: 10.1109/LGRS.2020.3023706, 2022.

[15] J. Mendel, White-noise estimators for seismic data processing in oil exploration, *IEEE Transactions on Automatic Control*, vol.22, no.5, pp.694-706, doi: 10.1109/TAC.1977.1101597, 1977.

[16] C. Pinnegar and L. Mansinha, The S-transform with windows of arbitrary and varying shape, *Geophysics*, vol.68, no.1, pp.381-385, 2003.

[17] B. Boashash and M. Mesbah, Signal enhancement by time-frequency peak filtering, *IEEE Transactions on Signal Processing*, vol.52, no.4, pp.929-937, 2004.

[18] N. Wu, Y. Li and B. Yang, Noise attenuation for 2-D seismic data by radial-trace time-frequency peak filtering, *IEEE Geoscience and Remote Sensing Letters*, vol.8, no.5, pp.874-878, 2011.

[19] X. Dong, Y. Li and B. Yang, Desert low-frequency noise suppression by using adaptive DnCNNs based on the determination of high-order statistic, *Geophysical Journal International*, vol.219, no.2, pp.1281-1299, 2019.

[20] Y. Deng, W. Wang, C. Qian et al., Boundary-processing-technique in EMD method and Hilbert transform, *Chin. Sci. Bull.*, vol.46, pp.954-960, https://doi.org/10.1007/BF02900475, 2001.
[21] Y. Kopsinis and S. McLaughlin, Development of EMD-based denoising methods inspired by wavelet thresholding, *IEEE Transactions on Signal Processing*, vol.57, no.4, pp.1351-1362, 2009.
[22] J. Han and V. B. Mirko, Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding, *Geophysics*, vol.80, no.6, pp.KS69-KS80, doi: https://doi.org/10.1190/geo2014-0423.1, 2015.
[23] Y. Liu, G. Yang, M. Li et al., Variational mode decomposition denoising combined the detrended fluctuation analysis, *Signal Processing*, vol.125, pp.349-364, 2016.
[24] J. Sanchez, O. Alegria, M. Rodriguez, J. Abeyro, J. Almaraz and A. Gonzalez, Detection of ULF geomagnetic anomalies associated to seismic activity using EMD method and fractal dimension theory, *IEEE Latin America Transactions*, vol.15, no.2, pp.197-205, 2017.
[25] G. Ren and Z. Liu, An improved EMD adaptive denoising and feature extraction algorithm, *2019 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp.1-4, 2019.
[26] A. Lewis and G. Knowles, Image compression using the 2-D wavelet transform, *IEEE Transactions on Image Processing*, vol.1, no.2, pp.244-250, doi: 10.1109/83.136601, 1992.
[27] L. Li and Y. Si, Remote sensing image enhancement based on nonsubsampled shearlet transform and local Laplacian filter, *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, pp.415-418, doi: 10.1109/ICIVC.2018.8492906, 2018.
[28] Z. Geng, X. Wu, S. Fomel et al., Relative time seislet transform, *Geophysics*, vol.85, no.2, pp.V223-V232, 2020.
[29] E. Uslu and S. Albayrak, Curvelet-based synthetic aperture radar image classification, *IEEE Geoscience and Remote Sensing Letters*, vol.11, no.6, pp.1071-1075, doi: 10.1109/LGRS.2013.2286089, 2014.
[30] Z. Fen and X. Kai, Regularization over-complete dictionary learning with application to image denoising, *IEEE International Conference on Intelligent Control, Automatic Detection and High-End Equipment*, pp.30-34, 2012.
[31] K. Frank and P. Richard, Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics*, vol.133, no.1, pp.97-126, 2006.
[32] S. Malini and R. S. Moni, Image denoising using multi-resolution principal component analysis, *2015 Global Conference on Communication Technologies (GCCT)*, pp.4-7, 2015.
[33] L. Wang, X. Wang and L. Liu, Application of anisotropic diffusion filtering in seismic data processing, *2011 International Conference on Computer Science and Service System (CSSS)*, pp.3046-3049, 2011.
[34] L. Yang, W. Chen, H. Wang and Y. Chen, Deep learning seismic random noise attenuation via improved residual convolutional neural network, *IEEE Transactions on Geoscience and Remote Sensing*, vol.59, no.9, pp.7968-7981, 2021.
[35] G. Yu, Z. Cai, Y. Chen, X. Wang et al., Borehole seismic survey using multimode optical fibers in a hybrid wireline, *Measurement*, vol.125, pp.694-703, 2018.
[36] B. Liu, S. Yang, Y. Ren, X. Xu, P. Jiang and Y. Chen, Deep-learning seismic full-waveform inversion for realistic structural models, *Geophysics*, vol.86, no.1, pp.R31-R44, 2021.
[37] F. Wang and S. Chen, Residual learning of deep convolutional neural network for seismic random noise attenuation, *IEEE Geoscience and Remote Sensing Letters*, vol.16, no.8, pp.1314-1318, 2019.
[38] H. Zhang, X. Yang and J. Ma, Can learning from natural image denoising be used for seismic data interpolation?, *Geophysics*, vol.85, no.4, pp.WA115-WA136, 2020.
[39] H. Zhang, Y. Zhang, S. Zhu and D. Xu, Deterministic convergence of complex mini-batch gradient learning algorithm for fully complex-valued neural networks, *Neurocomputing*, vol.407, pp.185-193, 2020.
[40] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, *IEEE Transactions on Image Processing*, vol.26, no.7, pp.3142-3155, 2017.
[41] Y. Zhao, Y. Li, X. Dong and B. Yang, Low-frequency noise suppression method based on improved DnCNN in desert seismic data, *IEEE Geoscience and Remote Sensing Letters*, vol.16, no.5, pp.811-815, 2019.
[42] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. Courville, Improved training of Wasserstein GANs, *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, Curran Associates Inc., Red Hook, NY, USA, pp.5769-5779, 2017.

[43] J. Zhang, Q. Qi, H. Zhang, Q. Du, Z. Guo and Y. Tian, Detection of bird's nest on transmission lines from aerial images based on deep learning model, *International Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1755-1768, 2022.
[44] T. Zhong, M. Cheng, X. Dong and N. Wu, Seismic random noise attenuation by applying multi-scale denoising convolutional neural network, *IEEE Transactions on Geoscience and Remote Sensing*, vol.60, 5905013, doi: 10.1109/TGRS.2021.3095922, 2022.
[45] X. Dong, T. Zhong and Y. Li, New suppression technology for low-frequency noise in desert region: The improved robust principal component analysis based on prediction of neural network, *IEEE Transactions on Geoscience and Remote Sensing*, vol.58, no.7, pp.4680-4690, doi: 10.1109/TGRS.2020.2966054, 2020.
[46] Y. Zhang, J. Wei, D. Xu et al., Batch gradient training method with smoothing group l0 regularization for feedfoward neural networks, *Neural Process. Lett.*, 2022.
[47] K. Hashimoto, T. Kamiya, G. Li, K. Yoneda and F. Tanaka, Automatic identification of tumor cells for circulating tumor cells by convolutional neural networks, *International Journal of Innovative Computing, Information and Control*, vol.19, no.1, pp.1-14, 2023.
[48] J. Souza, G. Santana, L. Batista, G. Oliveira, E. Roemers-Oliveira and M. Santos, CNN prediction enhancement by post-processing for hydrocarbon detection in seismic images, *IEEE Access*, vol.8, pp.120447-120455, 2020.
[49] P. Wang et al., Understanding convolution for semantic segmentation, *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.1451-1460, 2018.

Author Biography

Tie Zhong received the B.S., M.S., and Ph.D. degrees in Communication Engineering from Jilin University, Changchun, China, in 2006, 2008 and 2016, respectively. He is currently an Associate Professor in the Department of Communication Engineering, Northeast Electric Power University, Jilin, China. His research interests include modern signal processing techniques, weak signal process theory, and intelligent information processing.

Haoliang Chen received his bachelor's degree in Communication Engineering from Northeast Electric Power University, Jilin, China in 2019. He is studying for a master's degree in Northeast Electric Power University, Jilin, China. His main research interest is modern signal processing.

Shiqi Dong received the bachelor's degree in Applied Geophysics from Jilin University, Changchun, China, in 2016, where he is currently pursuing the Ph.D. degree in Geo-Exploration and Information Technology. He has been a Visiting Student with the Massachusetts Institute of Technology from August 2019 to October 2019. His current research interests include full waveform inversion and deep learning.

Jianpo Li (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the Department of Communication Engineering, Jilin University, China, in 2002, 2005, and 2008, respectively. In 2008, he joined the School of Computer Science, Northeast Electric Power University, where he is currently a Professor. He was a Visiting Scholar with New York University and the University of Ottawa, in 2013 and 2016, respectively. His research interests include wireless sensor networks and intelligent signal processing.